Workgroup: SFC
Internet-Draft: draft-ietf-sfc-nsh-tlv-14
Published: 30 March 2022
Intended Status: Standards Track
Expires: 1 October 2022
Authors: Yuehua. Wei, Ed. U. Elzur S. Majee
ZTE Corporation Intel Individual contributor
C. Pignataro D. Eastlake
Cisco Futurewei Technologies

Network Service Header Metadata Type 2 Variable-Length Context Headers

Abstract

Service Function Chaining (SFC) uses the Network Service Header (NSH) (RFC 8300) to steer and provide context Metadata (MD) with each packet. Such Metadata can be of various Types including MD Type 2 variable length context headers. This document specifies several such context headers that can be used within a service function path.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 1 October 2022.

Table of Contents

1. Introduction

The Network Service Header (NSH) [RFC8300] is the Service Function Chaining (SFC) encapsulation that supports the SFC architecture [RFC7665]. As such, the NSH provides following key elements:

1. Service Function Path (SFP) identification.

2. Indication of location within a Service Function Path.

3. Optional, per-packet metadata (fixed-length or variable-length).

[RFC8300] further defines two metadata formats (MD Types): 1 and 2. MD Type 1 defines the fixed-length, 16-octet long metadata, whereas MD Type 2 defines a variable-length context format for metadata. This document defines several common metadata context headers for use with NSH MD Type 2. These supplement the Subscriber Identity and Performance Policy MD Type 2 metadata context headers specified in [RFC8979].

This document does not address metadata usage, updating/chaining of metadata, or other SFP functions. Those topics are described in [RFC8300].

2. Conventions used in this document

2.1. Terminology

This document uses the terminology defined in the SFC Architecture [RFC7665] and the Network Service Header [RFC8300].

2.2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. NSH MD Type 2 format

An NSH is composed of a 4-octet Base Header, a 4-octet Service Path Header and optional Context Headers. The Base Header identifies the MD-Type in use:

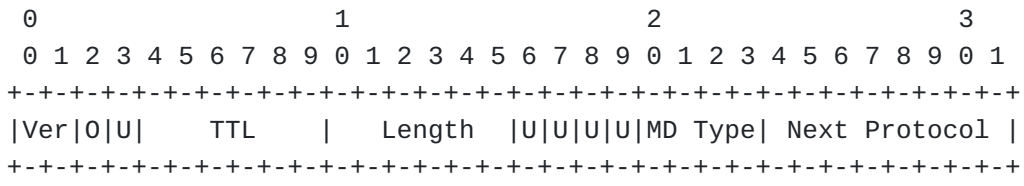

Figure 1: NSH Base Header

Please refer to NSH [RFC8300] for a detailed header description.

When the base header specifies MD Type = 0x2, zero or more Variable Length Context Headers MAY be added, immediately following the Service Path Header. Figure 2 below depicts the format of the Context Header as defined in Section 2.5.1 of [RFC8300].

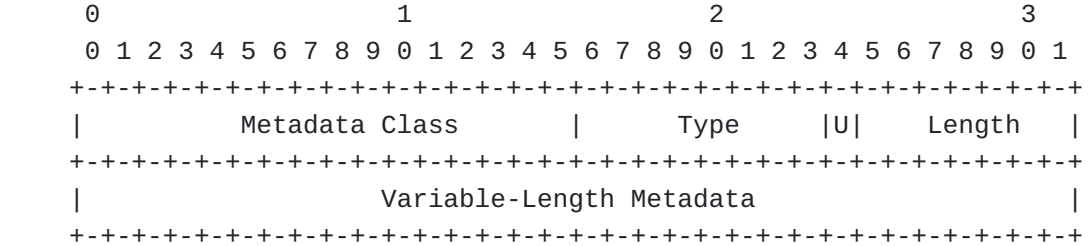

Figure 2: NSH Variable-Length Context Headers

4. NSH MD Type 2 Context Headers

[RFC8300] specifies Metadata Class 0x0000 as IETF Base NSH MD Class. In this document, metadata types are defined for the IETF Base NSH MD Class.

4.1. Forwarding Context

This metadata context carries a network forwarding context, used for segregation and forwarding scope. Forwarding context can take several forms depending on the network environment. For example, VXLAN/VXLAN-GPE VNID, VRF identification, or VLAN.

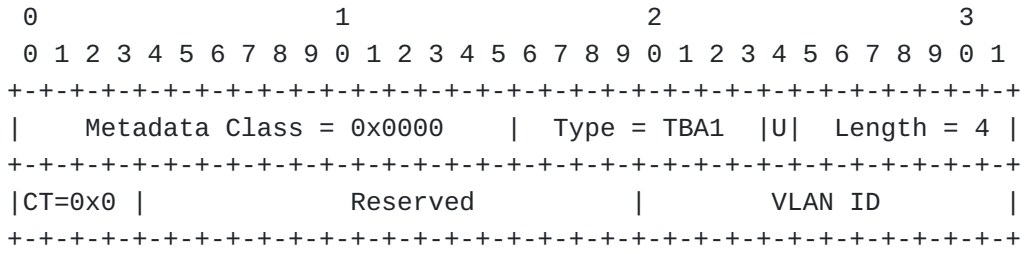

Figure 3: Forwarding Context - 1(VLAN)

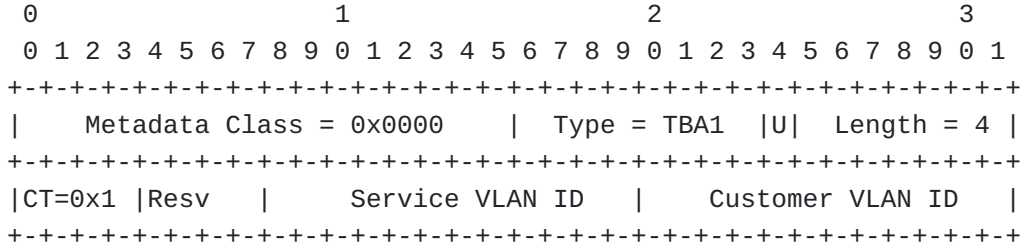

Figure 4: Forwarding Context - 2(QinQ)

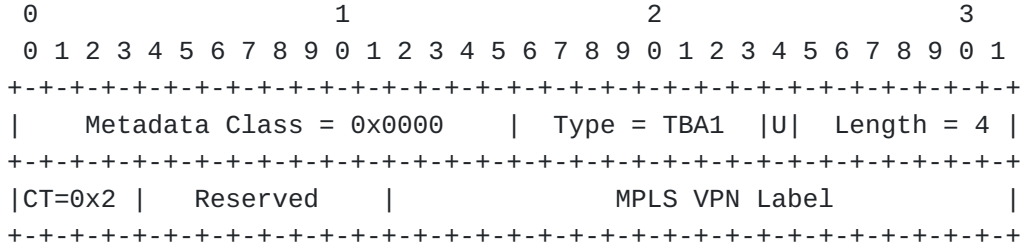

Figure 5: Forwarding Context - 3(MPLS VPN)

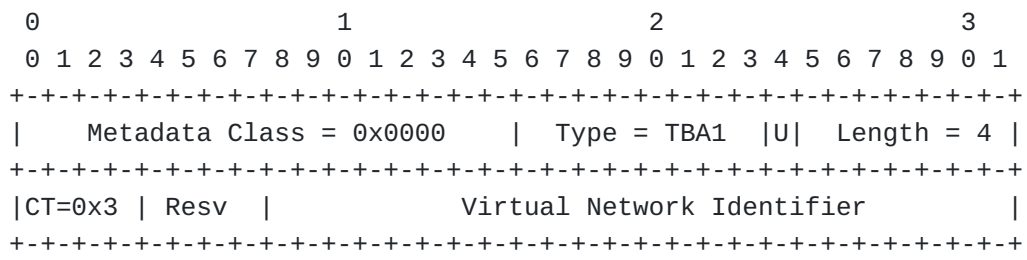

Figure 6: Forwarding Context - 4(VNI)

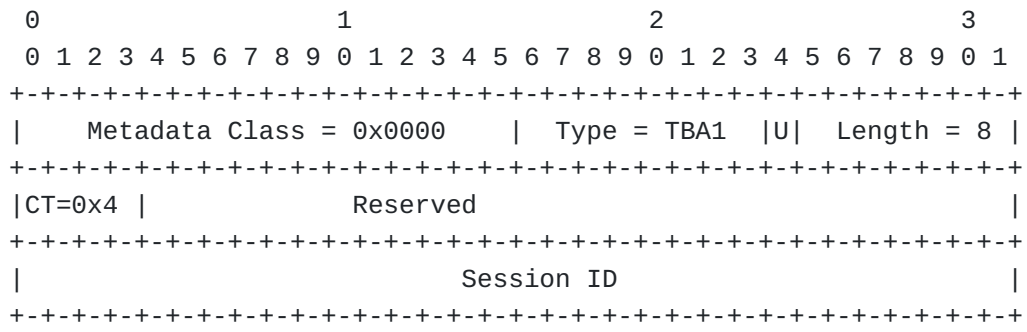

Figure 7: Forwarding Context - 5(Session ID)

where:

Context Type (CT) is four bits-long field that defines the length and the interpretation of the Forwarding Context field. Please see the IANA Considerations in Section 7.2. This document defines these CT values:

- 0x0 - 12 bits VLAN identifier [IEEE.802.1Q_2018]. See Figure 3.

- 0x1 - 24 bits double tagging identifiers. A service VLAN tag followed by a customer VLAN tag [IEEE.802.1Q_2018]. The two VLAN IDs are concatenated and appear in the same order that they appeared in the payload. See Figure 4.

- 0x2 - 20 bits MPLS VPN label([RFC3032])([RFC4364]). See Figure 5.

- 0x3 - 24 bits virtual network identifier (VNI)[RFC8926]. See Figure 6.

- 0x4 - 32 bits Session ID ([RFC3931]). This is called Key in GRE [RFC2890]. See Figure 7.

Reserved bits in the context fields MUST be sent as zero and ignored on receipt.

4.2. Tenant Identifier

Tenant identification is often used for segregation within a multi-tenant environment. Orchestration system-generated tenant IDs are an example of such data. This context header carries the value of the Tenant identifier. [OpenDaylight-VTN] Virtual Tenant Network (VTN) is an application that provides multi-tenant virtual network on an SDN controller.

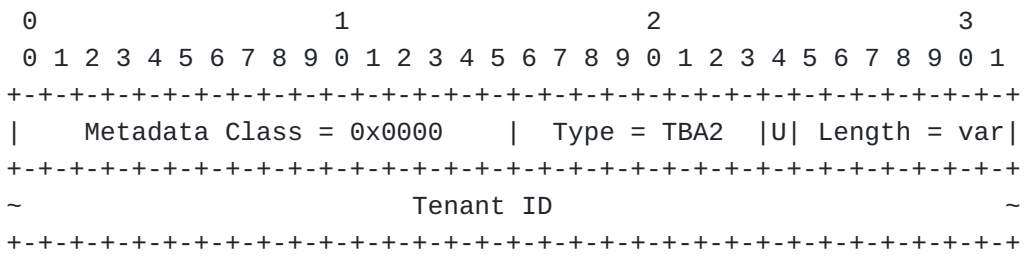

Figure 8: Tenant Identifier List

The fields are described as follows:

Length: Indicates the length of the Tenant ID in octets (see Section 2.5.1 of [RFC8300]).

Tenant ID: Represents an opaque value pointing to Orchestration system-generated tenant identifier. The structure and semantics of this field are specific to the operator's deployment across its operational domain, and are specified and assigned by an orchestration function. The specifics of that orchestration-based assignment are outside the scope of this document.

4.3. Ingress Network Node Information

This context header carries a Node ID of the ingress network node.

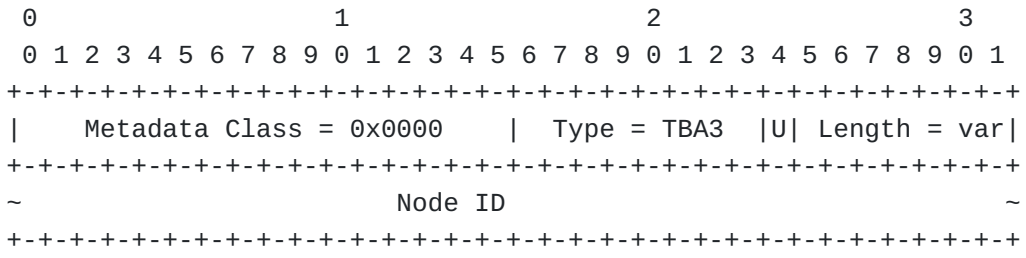

Figure 9: Ingress Network Node ID

The fields are described as follows:

Length: Indicates the length of the Node ID in octets (see Section 2.5.1 of [RFC8300]).

Node ID: Represents an opaque value of the ingress network node ID. The structure and semantics of this field are deployment specific. For example, Node ID may be a 4 octets IPv4 address Node ID, or a 16 octets IPv6 address Node ID, or a 6 octets MAC address, or 8 octets MAC address (EUI-64), etc.

4.4. Ingress Network Source Interface

This context identifies the ingress interface of the ingress network node. The l2vlan (135), l3ipvlan (136), ipForward (142), mpls (166) in [IANAifType] are examples of source interfaces.

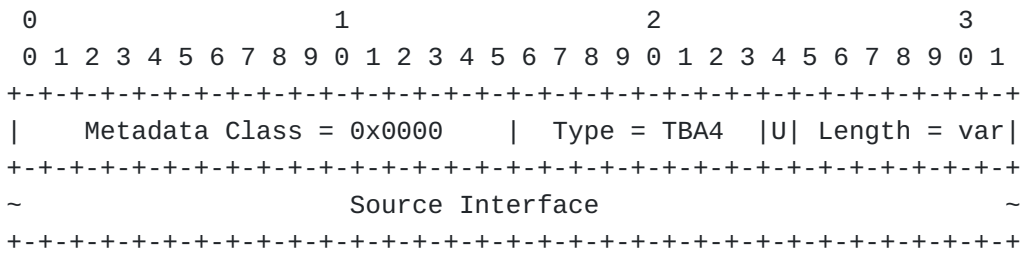

Figure 10: Ingress Network Source Interface

The fields are described as follows:

Length: Indicates the length of the Source Interface in octets (see Section 2.5.1 of [RFC8300]).

Source Interface: Represents an opaque value of identifier of the ingress interface of the ingress network node.

4.5. Flow ID

Flow ID provides a field in the NSH MD Type 2 to label packets belonging to the same flow. For example, [RFC8200] defined IPv6 Flow Label as Flow ID, [RFC6790] defined an entropy label which is generated based on flow information in the MPLS network is another example of Flow ID. Absence of this field, or a value of zero denotes that packets have not been labeled.

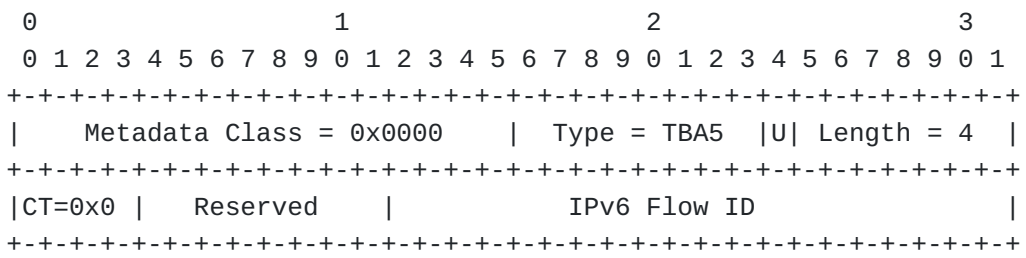

Figure 11: IPv6 Flow ID

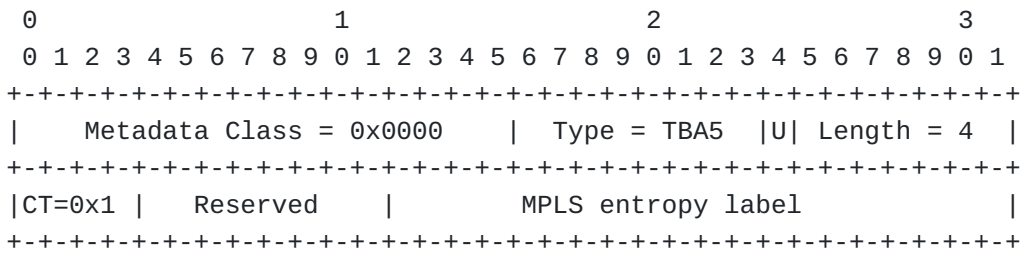

Figure 12: MPLS entropy label

The fields are described as follows:

Length: Indicates the length of the Flow ID in octets (see Section 2.5.1 of [RFC8300]). For example, IPv6 Flow Label in [RFC8200] is 20-bit long. An entropy label in the MPLS network in [RFC6790] is also 20-bit long.

Context Type (CT) is four bits-long field that defines the length and the interpretation of the Flow ID field. Please see the IANA Considerations in Section 7.3. This document defines these CT values:

- 0x0 - 20 bits IPv6 Flow Label in [RFC8200]. See Figure 11.

- 0x1 - 20 bits entropy label in the MPLS network in [RFC6790]. See Figure 12.

Reserved bits in the context fields MUST be sent as zero and ignored on receipt.

4.6. Source and/or Destination Groups

Intent-based systems can use this data to express the logical grouping of source and/or destination objects. [OpenStack] and [OpenDaylight] provide examples of such a system. Each is expressed as a 32-bit opaque object.

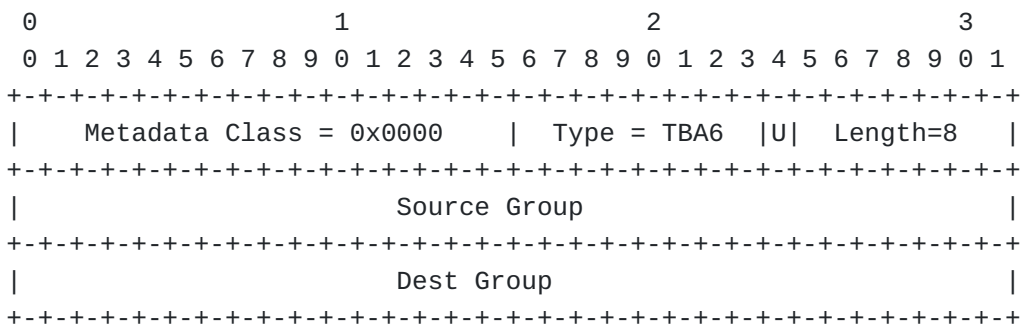

Figure 13: Source/Dest Groups

If there is no group information specified for the source group or dest group field, the field MUST be sent as zero and ignored on receipt.

4.7. Policy Identifier

Traffic handling policies are often referred to by a system-generated identifier, which is then used by the devices to look up the policy's content locally. For example, this identifier could be an index to an array, a lookup key, a database Id. The identifier allows enforcement agents or services to look up the content of their part of the policy.

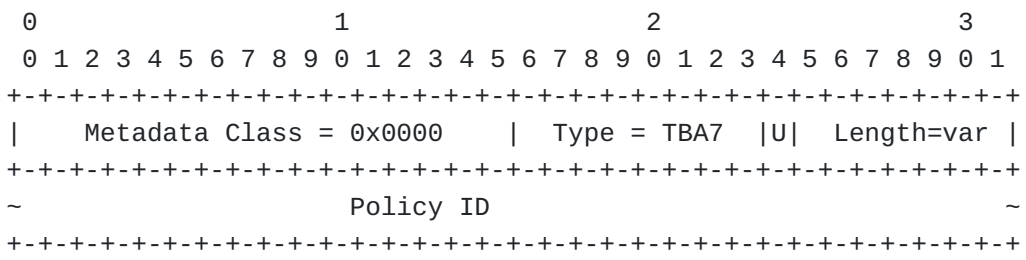

Figure 14: Policy ID

The fields are described as follows:

Length: Indicates the length of the Policy ID in octets (see Section 2.5.1 of [RFC8300]).

Policy ID: Represents an opaque value of the Policy ID.

This policy identifier is a general policy ID, essentially a key to allow Service Functions to know which policies to apply to packets. Those policies generally will not have much to do with performance, but rather with what specific treatment to apply. It may for example select a URL filter data set for a URL filter, or select a video transcoding policy in a transcoding SF. The Performance Policy

Identifier in [RFC8979] is described there as having very specific use, and for example says that fully controlled SFPs would not use it. The Policy ID in this document is for cases not covered by [RFC8979].

5. Security Considerations

A misbehaving node from within the SFC-enabled domain may alter the content of the Context Headers, which may lead to service disruption. Such an attack is not unique to the Context Headers defined in this document. Measures discussed in Section 8 of [RFC8300] describes the general security considerations for protecting NSH. [I-D.ietf-sfc-nsh-integrity] specifies methods of protecting the integrity of the NSH metadata. If the NSH includes the MAC Context Header, the authentication of the packet MUST be verified before using any data. If the verification fails, the receiver MUST stop processing the variable length context headers and notify an operator.

6. Acknowledgments

The authors would like to thank Paul Quinn, Behcet Sarikaya, Dirk von Hugo, Mohamed Boucadair, Gregory Mirsky, and Joel Halpern for providing invaluable concepts and content for this document.

7. IANA Considerations

7.1. MD Type 2 Context Types

IANA is requested to assign the following types (Table 1) from the "NSH IETF- Assigned Optional Variable-Length Metadata Types" registry available at [IANA-NSH-MD2].

Value	Description	Reference
TBA1	Forwarding Context	This document
TBA2	Tenant Identifier	This document
TBA3	Ingress Network NodeID	This document
TBA4	Ingress Network Interface	This document
TBA5	Flow ID	This document
TBA6	Source and/or Destination Groups	This document
TBA7	Policy Identifier	This document

Table 1: Type Values

7.2. Forwarding Context Types

IANA is requested to create a new sub-registry for "Forwarding Context" context types at [IANA-NSH-MD2] as follows:

The Registration Policy is IETF Review

Value	Forwarding Context Header Types	Reference
0x0	12-bit VLAN identifier	This document
0x1	24-bit double tagging identifiers	This document
0x2	20-bit MPLS VPN label	This document
0x3	24-bit virtual network identifier (VNI)	This document
0x4	32-bit Session ID	This document
0x5-0xE	Unassigned	
0xF	Reserved	This document

Table 2: Forwarding Context Types

7.3. Flow ID Context Types

IANA is requested to create a new sub-registry for "Flow ID Context" context types at [IANA-NSH-MD2] as follows:

The Registration Policy is IETF Review

Value	Flow ID Context Header Types	Reference
0x0	20-bit IPv6 Flow Label	This document
0x1	20-bit entropy label in the MPLS network	This document
0x2-0xE	Unassigned	
0xF	Reserved	This document

Table 3: Flow ID Context Types

8. References

8.1. Normative References

[I-D.ietf-sfc-nsh-integrity] Boucadair, M., Reddy, T., and D. Wing, "Integrity Protection for the Network Service Header (NSH) and Encryption of Sensitive Context Headers", Work in Progress, Internet-Draft, draft-ietf-sfc-nsh-integrity-09, 20 September 2021, <https://www.ietf.org/archive/id/draft-ietf-sfc-nsh-integrity-09.txt>.

[IANA-NSH-MD2] IANA, "NSH IETF-Assigned Optional Variable-Length Metadata Types", <https://www.iana.org/assignments/nsh/nsh.xhtml#optional-variable-length-metadata-types>.

[IEEE.802.1Q_2018] IEEE, "IEEE Standard for Local and Metropolitan Area Networks--Bridges and Bridged Networks", July 2018,

<http://ieeexplore.ieee.org/servlet/opac?punumber=8403925>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC3931] Lau, J., Ed., Townsley, M., Ed., and I. Goyret, Ed., "Layer Two Tunneling Protocol - Version 3 (L2TPv3)", RFC 3931, DOI 10.17487/RFC3931, March 2005, <https://www.rfc-editor.org/info/rfc3931>.

[RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <https://www.rfc-editor.org/info/rfc8174>.

[RFC8300] Quinn, P., Ed., Elzur, U., Ed., and C. Pignataro, Ed., "Network Service Header (NSH)", RFC 8300, DOI 10.17487/RFC8300, January 2018, <https://www.rfc-editor.org/info/rfc8300>.

8.2. Informative References

[IANAifType] IANA, "IANAifType", 2021, <https://www.iana.org/assignments/ianaiftype-mib/ianaiftype-mib>.

[OpenDaylight] OpenDaylight, "Group Based Policy", 2021, <https://docs.opendaylight.org/en/stable-fluorine/user-guide/group-based-policy-user-guide.html?highlight=group%20policy#>.

[OpenDaylight-VTN] OpenDaylight, "OpenDaylight VTN", 2021, <https://nexus.opendaylight.org/content/sites/site/org.opendaylight.docs/master/userguide/manuals/userguide/bk-user-guide/content/_vtn.html>.

[OpenStack] OpenStack, "Group Based Policy", 2021, <https://wiki.openstack.org/wiki/GroupBasedPolicy>.

[RFC2890] Dommety, G., "Key and Sequence Number Extensions to GRE", RFC 2890, DOI 10.17487/RFC2890, September 2000, <https://www.rfc-editor.org/info/rfc2890>.

[RFC3032] Rosen, E., Tappan, D., Fedorkow, G., Rekhter, Y., Farinacci, D., Li, T., and A. Conta, "MPLS Label Stack

Encoding", RFC 3032, DOI 10.17487/RFC3032, January 2001, <https://www.rfc-editor.org/info/rfc3032>.

[RFC4364] Rosen, E. and Y. Rekhter, "BGP/MPLS IP Virtual Private Networks (VPNs)", RFC 4364, DOI 10.17487/RFC4364, February 2006, <https://www.rfc-editor.org/info/rfc4364>.

[RFC6790] Kompella, K., Drake, J., Amante, S., Henderickx, W., and L. Yong, "The Use of Entropy Labels in MPLS Forwarding", RFC 6790, DOI 10.17487/RFC6790, November 2012, <https://www.rfc-editor.org/info/rfc6790>.

[RFC7665] Halpern, J., Ed. and C. Pignataro, Ed., "Service Function Chaining (SFC) Architecture", RFC 7665, DOI 10.17487/RFC7665, October 2015, <https://www.rfc-editor.org/info/rfc7665>.

[RFC8200] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", STD 86, RFC 8200, DOI 10.17487/RFC8200, July 2017, <https://www.rfc-editor.org/info/rfc8200>.

[RFC8926] Gross, J., Ed., Ganga, I., Ed., and T. Sridhar, Ed., "Geneve: Generic Network Virtualization Encapsulation", RFC 8926, DOI 10.17487/RFC8926, November 2020, <https://www.rfc-editor.org/info/rfc8926>.

[RFC8979] Sarikaya, B., von Hugo, D., and M. Boucadair, "Subscriber and Performance Policy Identifier Context Headers in the Network Service Header (NSH)", RFC 8979, DOI 10.17487/RFC8979, February 2021, <https://www.rfc-editor.org/info/rfc8979>.

Authors' Addresses

Yuehua Wei (editor)
ZTE Corporation
No.50, Software Avenue
Nanjing
210012
China

Email: wei.yuehua@zte.com.cn

Uri Elzur
Intel

Email: uri.elzur@intel.com

Sumandra Majee

Individual contributor

Email: Sum.majee@gmail.com

Carlos Pignataro
Cisco

Email: cpignata@cisco.com

Donald E. Eastlake
Futurewei Technologies

Email: d3e3e3@gmail.com